Medicare Enrollment and Claims Data Analytic Issues

General Notices to Users

This document provides additional information about Medicare utilization & enrollment data. Users also should refer to the Matching Methodology report and the Description of NCHS-CMS Linkage document. This document is not an exhaustive or systematic review of the analytic issues researchers may encounter while using the NCHS-Medicare Linked Data files and it will be updated as additional analytic issues are identified and brought to the attention of the NCHS Research Data Center or Data Linkage Unit. Users are encouraged to visit the ResDAC website www.resdac.umn.edu for more information on Medicare data.

The advantages of Medicare data is that it is population-based, not subject to recall bias, and has been linked to NCHS population health surveys. However because Medicare data were collected for the purpose of making healthcare payments, and not for research, there are limitations to the data that researchers should consider when constructing their samples and conducting analyses.

1. Denominator file
All applications to the RDC should include a request for the Denominator file for the years that the researcher is examining claims data. The Denominator Record contains basic demographic and enrollment information about each beneficiary enrolled during each calendar year and is needed to help construct your sample, particularly to identify Medicare beneficiaries enrolled in a managed care plan (see below).

2. Exclusion of claims paid by a source other than Medicare (e.g., managed care organizations or HMOs)
CMS generally does not receive claims data for Medicare beneficiaries who enroll in managed care plans (including private fee-for-service plans paid on a capitation basis). During the time covered by the linked database, enrollment in managed care increased from approximately 6% of beneficiaries in 1991 to 17% in 1999. A large number of managed care plans withdrew from Medicare beginning in 1999, resulting in a decrease in enrollment in 2000 to 16% of beneficiaries. In general, studies based on analysis of claims data should exclude managed care enrollees from their beneficiary samples. A technical brief about health care utilization information for Medicare beneficiaries enrolled in a managed care plan and how to exclude them from your sample can be found at http://www.resdac.umn.edu/Tools/TBs/TN-009.asp.

3. Services not covered
Although Medicare provides coverage for a wide range of services, there are health care services not covered by Medicare as well as a number of cost sharing requirements for Medicare beneficiaries. Examples of services not covered include routine physical exams, long-term care, and some cancer screening procedures. These gaps in coverage and required cost-sharing mean that there are no claims records for these services or for

certain time periods. You may find more information on what is not covered by Medicare at www.cms.hhs.gov or www.aarp.org/health/medicare.

In addition, Medicare data contains little information on prescription drugs for years prior to 2006. Medicare enrollment and utilization data linked to NCHS survey data is only available only for the years 1991-2000. Prescription drug information that is paid for by Medicare and available for the data years 1991-2000 includes:

- Medication given in an inpatient/hospice/SNF setting - but note that the specific medicines dispensed are rarely coded.
- The administration of chemotherapy if chemotherapy is administered intravenously (IV), chemotherapy is administered orally as a substitute for a medication that could be administered IV, or the oral chemotherapeutic agent is a drug that breaks down to a compound comparable to a chemotherapeutic agent administered IV.
- Medicare does not pay for chemotherapeutic agents that are administered exclusively in an oral form (e.g., Tamoxifen) and prior to 2006 most outpatient prescription drugs were not covered by Medicare.

4. Gaps in coverage

Medicare enrollment and utilization data is available only for the years 1991-2000. Several of the surveys linked to the Medicare data, such as NHEFS, NHANES II, and NHANES III have gaps of several years between the end of the study period and the beginning of the Medicare data.

Additional analytic issues specific to each of the Medicare files are described below.

Denominator

The Denominator file provides data on all Medicare beneficiaries enrolled and/or entitled to Medicare benefits in a given year. Monthly information on the enrollment status of linked Medicare beneficiaries including third party payer information and Group Health Plan (GHP) enrollment information is provided. The Denominator file is fixed length and contains one record per person.

Date of death information obtained by CMS is available on the Denominator File. CMS collects death information for each calendar year through the first three months of the following calendar year. Deaths to Medicare eligible beneficiaries occurring in the first quarter of the year will be recorded on that year's denominator file but may also be recorded on the previous year's denominator file. For example, a CMS recorded death occurring on 02/01/1995 will have a date recorded in variable DOD 'Date of Death' on the 1995 Denominator file and may also have that same date recorded in variable DOD 'Date of Death' on the 1994 Denominator file. In addition, death information is occasionally mis-reported to CMS but included on the yearly Denominator file. This erroneous information is not corrected by CMS; however, these cases can be identified as they continue to be eligible for Medicare benefits in later years or they have new death

information recorded in a later Denominator file. Analysts should use extra caution in analyzing Medicare death information to insure that deaths are not over-counted.

Linked Mortality data files developed from probabilistic matches to the National Death Index, death certificates, or longitudinal survey re-contacts are also available through the NCHS RDC for the 1994-1998 NHIS, NHANES I, NHANES II, NHANES III, and LSOA II. No attempt has been made to reconcile inconsistent death information from CMS and other sources. RDC research proposals that intend to analyze mortality and morbidity outcomes should utilize both sources of data.

Documentation for the Denominator File is available in PDF format. The variable names used in this data file come from the suggested SAS alias variable name provided by CMS in the Denominator file documentation.

Medicare Provider Analysis and Review File (MedPAR) Hospital Stay

The MedPAR Hospital Stay file contains inpatient hospitalization final action claim records. All Medicare Part A short and long stay hospitalization claims for each calendar year are included on the MedPAR Hospital Stay file. Each MedPAR Hospital Stay claim record includes up to 10 ICD-9 diagnoses and 6 ICD-9 procedures associated with each hospital stay. Claim record inclusion on the MedPAR Hospital Stay file is based on year of discharge.

Claims for hospital stays starting in one calendar year and continuing past the end of the calendar year are not provided on the MedPAR file until the year of discharge. To determine if a claim record is for a long stay or short stay hospitalization use variable 'MEDPAR_SS_LS_SNF_IND_CD' - Short Stay/Long Stay/SNF Indicator' which is coded S for short stay or L for long stay.

In most cases, a single MedPAR Hospital Stay claim record reflects a summary of all care provided during the hospital stay. However, if the stay is long, there may be more than one claim record per hospital stay.

There are several fields on the MedPAR Hospital Stay files that CMS considers unreliable:

- source of admission
  MEDPAR_SRC_IP_ADMSN_CD

- discharge destination
  MEDPAR_DSCHRG_DSTNTN_CD

- group health organization payment code
  MEDPAR_GHO_PD_CD

Documentation for the MedPAR Hospital Stay File is available in PDF format. The variable names used in this data file come from the suggested standard alias variable names provided by CMS in the MedPAR file documentation.

Medicare Provider Analysis and Review File (MedPAR) Skilled Nursing Facility (SNF)

The MedPAR Skilled Nursing Facility file contains skilled nursing facility final action claim records. Skilled Nursing Facility (SNF) claims for each calendar year are provided on the MedPAR SNF file. Each MedPAR SNF claim record includes up to 10 ICD-9 diagnoses and 6 ICD-9 procedures provided association with a SNF stay. Inclusion in the MedPAR SNF is based on year of admission into the facility.

In most cases, a single MedPAR SNF claim record reflects a summary of all care provided during an institutional stay. However, if the stay is long, there may be more than one claim record per single facility stay. This occurs most frequently for stays in SNFs which often span several months. In many cases, SNF records have no discharge date as persons remain in institutions beyond the period of Medicare coverage for that year.

There are several fields on the MedPAR SNF files that CMS are considers unreliable:

- source of admission
  SRC_IP_ADMSN_CD

- discharge destination
  DSCHRG_DSTNTN_CD

- group health organization payment code
  GHO_PD_CD

Documentation for the MedPAR Skilled Nursing Facility File is available in PDF format. The variable names used in this data file come from the suggested standard alias variable names minus the beginning text string “MEDPAR_” in the MedPAR file documentation provided by CMS.

Carrier

The Carrier file (formerly the Physician/Supplier Part B file) contains final action claims data submitted by non-institutional providers. The data is largely made up of physician claim records, although the file also includes claims from other non-institutional providers such as physician assistants, clinical social workers, nurse practitioners, independent clinical laboratories, ambulance providers, and stand-alone ambulatory surgical centers.

The claims are processed by private carriers working under contract to CMS. Each carrier claim must include a Health Care Procedure Classification Code (HCPCS) to describe the nature of the billed service. The HCPCS are composed primarily of CPT-4 codes developed by the American Medical Association, with additional codes specific to CMS. Each HCPCS code on the carrier claim must be accompanied by an ICD-9 diagnosis code, providing a reason for the service. In addition, each record includes the date of service and reimbursement amount. Due to the large number of carrier claim variables, CMS provides the Carrier data in variable length files. There can be multiple carrier claims per person on a file. The Carrier files are provided in the CMS Standard Analytic File format.

Important Information regarding the Carrier file:

The Carrier file includes records for non-institutional claims; however this does not mean that they are outpatient claims. Providers, such as physicians, can bill for services provided in the office, hospital, or other sites. The variable PLCSRVC 'Line Place of Service Code' tells where the service was provided.

The Carrier file contains Durable Medical Equipment (DME) claims processed by carriers who also process physician claims. The DME line items on the Carrier file can be identified by Claim Type Code (CLM_TYPE) equal to '72'. DME Claims processed through DME regional carriers are found on the DMERC files not on the carrier file. The DME claims on the Carrier file are for separate services than those on the DMERC file. There is no overlap between the DME claims on the Carrier and the DMERC files. See the section on Durable Medical Equipment (DMERC) below for additional information on DME regional carrier claims.

There are two pairs of date fields on the Carrier file. The variables FROM_DT 'Claim From Date' and THRU_DT 'Claim Through Date" generally cover a period of service (but not always a single date of service), while the variables EXPNSDT1 'Line First Expense Date' and EXPNSDT2 'Line Last Expense Date' represent the specific day of service.

For every billed procedure (using a HCPCS code), there should be a corresponding ICD-9 diagnosis code (LINEDGNS) that provides the reason for the billed service. In the case of lab tests, the diagnosis will often be XXOOO because the outside lab has no information from the physician about the reason for the test. In addition, the carrier file contains space for up to 4 diagnoses, DGNS_CD1 – DGNS_CD4. These are not necessarily linked with any of the billed procedures and may reflect co-existing health conditions. The accuracy of the diagnoses on the carrier data has not been determined by CMS.

Selected services may not appear in the carrier claims, although they may have been received by the claimant. For example, CMS pays physicians a fixed amount for surgeries. This practice is called bundling. As part of bundling, CMS expects that certain care will be included in the payment amount, such as the first one or two office visits

following surgery or a biopsy just before surgery. Bundled services will not appear in the physician data. How the rules on bundling are interpreted vary by carrier (physician).

Documentation for the Carrier SAF is available in PDF format. The variable names used in this data file come from the suggested SAS alias variable name provided by CMS in the Carrier SAF documentation.

Outpatient

The Outpatient file contains Medicare Part B final action claims from institutional outpatient providers for each calendar year. Hospital outpatient departments, rural health clinics, renal dialysis facilities, outpatient rehabilitation facilities, comprehensive outpatient rehabilitation facilities, and community mental health centers are examples of institutional outpatient providers. Same day surgeries performed in a hospital will be in the Outpatient file. However claims for surgeries performed in freestanding surgical centers appear in the Carrier file, not in the Outpatient file.

Some of the information provided in the Outpatient file includes diagnosis and procedure codes, dates of service, reimbursement amounts, revenue center codes, and some demographic information (such as date of birth, race, and sex). The Outpatient file contains data fields for 10 ICD-9 diagnosis and 6 procedure codes, but the reporting of these codes is sporadic. Services provided can be obtained from the HCPCS codes (HCPSCD01-HCPSCD45) and from the revenue centers (REV_CNTR). Definitions for revenue center codes can be found in the file documentation for the Outpatient file. There can be multiple outpatient claims records per person on the Outpatient files. The Outpatient files are provided in the CMS Standard Analytic File format.

Documentation for the Outpatient SAF is available in PDF format. The variable names used in this data file come from the suggested SAS alias variable name provided by CMS in the Outpatient SAF documentation.

Home Health Agency (HHA)

The Home Health Agency file contains final action claims for home health services. Some of the information contained in this file includes the number of visits, type of visit (skilled-nursing care, home health aides, physical therapy, speech therapy, occupational therapy, and medical social services), diagnosis (10 ICD-9 diagnosis codes), dates of visits, reimbursement amount. An HHA claim may cover services provided over a period of time, not a single day. There can be multiple HHA claims records per person on the HHA files. The Home Health Agency files are provided in the CMS Standard Analytic File format.

Documentation for the Home Health Agency SAF is available in PDF format. The variable names used in this data file come from the suggested SAS alias variable names provided by CMS in the Home Health Agency SAF documentation.

Hospice

The Hospice file contains final action claims data submitted by Hospice providers. The data contained in this file include the type of hospice care received (e.g., routine home care, inpatient respite care), The Hospice file contains data fields for 10 ICD-9 diagnosis and 6 procedure codes, dates of service, reimbursement amount, and some demographic information (such as date of birth, race, and sex). There can be multiple Hospice claims records per person on the Hospice file. The Hospice files are provided in the CMS Standard Analytic File format.

Documentation for the Hospice SAF is available in PDF format. The variable names used in this data file come from the suggested SAS alias variable names provided by CMS in the Hospice SAF documentation.

Durable Medical Equipment (DMERC)

The DMERC contains final action claims data submitted by Durable Medical Equipment (DME) regional carriers. Durable Medical Equipment can be billed through either a) the Carriers who also process physician claims or b) the DME Regional Carriers (DMERC's) who process only DME claims. Each year CMS distributes a jurisdiction list, available on the CMS website, which specifies whether a Carrier or a DMERC can process a claim for a particular service. Often, both Carriers and DMERCs are allowed to process and pay a DME claims service depending on whether or not the DME was provided "incident to the physician's service".

Some of the information contained in the DMERC includes diagnosis (10 ICD-9 diagnosis codes), service type codes, dates of service, and reimbursement amount. There can be multiple DME claim records per person on the DMERC file.

Important Note regarding DME claims data:

DME claims processed by suppliers who also process physician claims are only included on the Carrier file. These claims can be identified by Claim Type Code (CLM_TYPE) equal to '72' on the Carrier file. DME claims processed by regional carriers are only included on the DMERC file. Researchers should examine both the Carrier file and the DMERC file to obtain information about all DME claims. For years 1993-2000 approximately 90% of DME claims data are found on the DMERC file. However, for years 1991 and 1992 nearly 100% of the DME claims data are found on the Carrier file.

Documentation for Durable Medical Equipment is available in PDF format. The variable names used in this data file come from the suggested SAS alias variable names provided by CMS in the Durable Medical Equipment file documentation.

Acknowledgments

Information about the Medicare enrollment and utilization files was compiled from the following sources:

Centers for Medicare & Medicaid Services (CMS)
www.cms.hhs.gov

Research Data Assistance Center (ResDAC)
www.resdac.umn.edu

National Cancer Institute SEER-Medicare Linked Database
http://healthservices.cancer.gov/seermedicare/

Date Last Updated: April 26, 2006